A new factorization of mechanical words

Bruno Gaujal, Emmanuel Hyon

INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET EN AUTOMATIQUE

A new factorization of mechanical words

Bruno Gaujal — Emmanuel Hyon

N° 5175

Avril 2004

THÈME 1

Rapport de recherche

ISSN 0249-6399 ISRN INRIA/RR--5175--FR+ENG

INRIA RHÔNE-ALPES

A new factorization of mechanical words

Bruno Gaujal* , Emmanuel Hyon†

Thème 1 — Réseaux et systèmes
Projet TRIO

Rapport de recherche n° 5175 — Avril 2004 — 15 pages

Abstract: In this paper, we introduce a morphism on Sturmian words which is tightly related to the coefficients of a particular continued fraction the *ceiled continued fraction*. This morphism will be applied to factorize periodic Sturmian words called *Christoffel* words, as well as, to characterize and construct discrete geometric objects called *cellular line*.

Key-words: Ceiled continued fraction, Sturmian morphism, Christoffel words, Cellular lines.

* INRIA LIP, ENS Lyon, 46 allée d'Italie, F-69364 Lyon, France. E-mail Bruno.Gaujal@ens-lyon.fr
† LORIA, 615 route du jardin botanique, BP 101, F-54606 Villers-les-Nancy, France. E-mail: Emmanuel.Hyon@loria.fr.

Unité de recherche INRIA Rhône-Alpes
655, avenue de l'Europe, 38334 Montbonnot Saint Ismier (France)
Téléphone : +33 4 76 61 52 00 — Télécopie +33 4 76 61 52 52

Une nouvelle factorisation des mots mécaniques

Résumé : Dans ces travaux, nous introduisons un morphisme sur les mots de Sturm qui est intimement relié aux coefficients d'une fraction continue particulière : la fraction continue *ceiled*. Ce morphisme va être appliqué pour factoriser des "mots Sturmiens périodiques" appelés mots de *Christoffel*, tout comme pour construire et caractériser des objets de géométrie discrète : les lignes cellulaires.

Mots-clés : Fractions continues, morphismes sturmiens, lignes cellulaires, mots de Christoffel.

1 Introduction

Sturmian words are infinite binary words formally introduced in [16]. These words have been studied extensively because they appear in different fields such as physic [1], number theory [2], or combinatorics [14]. In particular, the factorization of aperiodic Sturmian words have been investigated in numerous papers (see [4] and references therein) and are generally done by multiple iterations of morphisms on words. In this paper, we present a new morphism which, combined with coefficients of a special expansion in continued fraction, can be applied to get a new factorization of mechanical words (also known as bracket sequences, Beatty sequences). Furthermore, this morphism has several nice properties: the number of iterations before convergence differs from this one obtained with the use of classical morphisms (it can be divided by two). Also it provides alternative and simple proofs of convergence theorems (such as Theorem 22). This morphism has also been used in optimal control [7, 8, 17]. Finally, this morphism provides a mean to construct discrete geometric objects, called *cellular lines*, with the help of mechanical words.

This paper is structured as follows : in section 2 we present preliminaries of mechanical words and continued fractions. The section 3 is devoted to the introduction of the morphism and its application to the factorization of mechanical words with ceiled continued fractions, while in section 4 we present the dual result by the factorization of mechanical words with a morphism derived to this one presented in the previous section. The last section, section 5, is dedicated to the presentation of several algorithmic applications of the morphism previously presented. In section 5.1, we will see how to check if a word is mechanical, while in section 5.2 we will show how to construct *cellular lines* from mechanical words.

2 Preliminaries

Here is a brief overview of the theory of words (finite as well as infinite). Most of the definitions and conventions in the following comes from [14] and [15].

Let $\mathcal{A} = \{0,1\}$ be the alphabet. The free monoid $\mathcal{A}^*$ is the set of the finite *words* on $\mathcal{A}$. An infinite word is an element of $\mathcal{A}^{\mathbb{N}}$. The set of finite and infinite words is denoted by $\mathcal{A}^\infty = \mathcal{A}^* \cup \mathcal{A}^{\mathbb{N}}$.

Let w be a word. The word f is called a *factor* of w if it exists two words (potentially empty) x and y such that $w = xfy$.

Hence a word w is factorized in $(f_1, f_2, \dots, f_n)$ if it exists a decomposition of w which only uses the factors $(f_1, f_2, \dots, f_n)$, but it may happen that this decomposition is not unique. A set of finite words is a code (resp. an ω-code) if for any word of $\mathcal{A}^*$ (resp. $\mathcal{A}^{\mathbb{N}}$) it exists at most one factorization using the words of this set.

Proposition 1 ([15]).
i) Any couple of words $\{x,y\}$ is a code except if it exists a word m and two integer number i and j, such that $x = m^i$ and $y = m^j$.
ii) Any code composed of two elements is an ω-code.

We denote by $\mathcal{S}$ the *shift* of a word. This is a morphism from $\mathcal{A}^{\mathbb{N}}$ onto $\mathcal{A}^{\mathbb{N}}$ such that if w is a word which nth letter is denoted by $w(n)$ then we have

$$\mathcal{S}(w(0)w(1)w(2)\dots) = w(1)w(2)w(3)\dots.$$

The composition of m shifts will be denoted by $\mathcal{S}_m$ and

$$\mathcal{S}_m(w(0)w(1)\dots) = w(m)w(m+1)\dots.$$

The *slope* of a finite nonempty word w is the number :

$$\mathbf{s}(w) = \frac{|w|_1}{|w|},$$

where $|w|_1$ is the number of letters equal to one in w and $|w|$ is the length of w.

Let $w_{[n]}, \forall n \geq 1$ be the prefix of length n of an infinite word w. If the sequence $\mathbf{s}(w_{[n]})$ converges when $n \to \infty$, the limit is called the slope of w.

A finite or infinite word w is *balanced* if for any factors x and y such that $|x| = |y|$ we have $| |x|_1 = |y|_1| \leq 1$.

2.1 Mechanical words

Infinite aperiodic mechanical words have been exhaustively studied in the past [15] mainly since they are *Sturmian* words. Finite (or periodic) mechanical words are called *Christoffel* words.

For a real number x, $\lfloor x \rfloor$ (resp. $\lceil x \rceil$) is the largest (resp. smallest) integer smaller (resp. larger) than x.

Definition 2 (Mechanical words). *The upper mechanical word with slope $0 \leq \alpha \leq 1$ and intercept θ is the infinite word $\overline{m}_\alpha^\theta$ where the nth letter $n \geq 0$, is :*

$$\overline{m}_\alpha^\theta(n) = \lceil (n+1) \times \alpha + \theta \rceil - \lceil n \times \alpha + \theta \rceil. \tag{1}$$

The lower mechanical word with slope α and intercept θ is the infinite word $\underline{m}_\alpha^\theta$ with:

$$\underline{m}_\alpha^\theta(n) = \lfloor (n+1) \times \alpha + \theta \rfloor - \lfloor n \times \alpha + \theta \rfloor. \tag{2}$$

The characteristic word of slope α and intercept θ is the infinite word c_α^θ with:

$$c_\alpha^\theta(n) = \lfloor (n+2) \times \alpha + \theta \rfloor - \lfloor (n+1) \times \alpha + \theta \rfloor.$$

In the following, $\overline{m}_\alpha$, $\underline{m}_\alpha$ will denote respectively the upper and the lower mechanical words of slope α and of intercept 0 and c_α is the characteristic word of intercept 0 of slope α.

Lemma 3 (Properties of mechanical words).
i) If α is a rational number ($\alpha = \frac{p}{q}$), then $\overline{m}_\alpha^\theta, \underline{m}_\alpha^\theta, c_\alpha^\theta$ are all periodic of period q. If α is an irrational number, then $\overline{m}_\alpha^\theta, \underline{m}_\alpha^\theta, c_\alpha^\theta$ are all aperiodic.
ii) Let α be a real number, then the maximum number of consecutive 0 in $\overline{m}_\alpha$ is $\lceil \alpha^{-1} \rceil - 1$ and the minimum number of consecutive 0 in $\overline{m}_\alpha$ is $\lfloor \alpha^{-1} \rfloor - 1$.

Proof. The proofs can be found in [15]. □

In the following, by a slight abuse of notation when an infinite word w is periodic we also denote by w its shortest period. The constraint $0 \leq \alpha \leq 1$ in Definition 2 can be relaxed. Indeed, if we assume that $k \leq \alpha < k+1$ with $k \in \mathbb{N}$, by the use of Formulas (2) and (1), we get that the words $\overline{m}_\alpha^\theta$, $\underline{m}_\alpha^\theta$ and c_α^θ are the mechanical and characteristic words over the alphabet $\{k, k+1\}$ which can be replaced by $\{0, 1\}$.

2.2 Classical Continued Fraction

Definition 4. *Let s be a real number, the computation of the expansion in the classical continued fraction of s is given by*

$$\left\{ \begin{array}{ccc} s = a_0 + s_0 & ; & a_0 = \lfloor s \rfloor \\ s_n = \frac{1}{a_{n+1} + s_{n+1}} & ; & a_{n+1} = \lfloor (s_n)^{-1} \rfloor \quad ; \forall n \geq 0 \end{array} \right\}. \tag{3}$$

When s is a rational number the expansion stops after a finite number of steps: there exists an integer number n such that $s_n = 1/a_{n+1}$. An infinite simple continued fraction is denoted by $[a_0, a_1, a_2, \ldots, a_n, \ldots]$. A finite continued fraction is denoted by $[a_0, a_1, \ldots, a_n]$, while a partial continued fraction of order n of the number s is denoted by $[a_0, \ldots, a_n + s_n]$.

Definition 5 (Convergent). *Let s be a real number which classical continued fraction is $[a_0, a_1, \ldots, a_n, \ldots]$. We define the rational number $\rho_n(s)$ as the classical convergent of order n of s by keeping only the n first coefficients of the continued fraction. We have $\rho_n(s) = [a_0, \ldots, a_n]$.*

It is well known that convergents form a sequence of rational numbers that converge to s as fast as possible, in the sense that no rational number with a denominator smaller that ρ_n is closer to s than ρ_n, [13].

2.3 Ceiled continued fractions

Ceiled continued fraction have be connected to the numbering of combinatoric objects see [11]. An example for the numbering of trajectories in lattice with ceiled continued fraction can be found [6].

Definition 6. *Let s be a real number such that* $0 < s < 1$*. The expansion in ceiled continued fraction of s is given by :*

$$\left\{ \begin{array}{lcl} s = 0 + \dfrac{1}{l_1 - s_1} & , & l_1 = \lceil s^{-1} \rceil \\ s_n = \dfrac{1}{l_{n+1} - s_{n+1}} & , & l_{n+1} = \lceil (s_n)^{-1} \rceil \quad , \forall n \geq 1 \end{array} \right\}. \tag{4}$$

If s is a rational number, the process stops after a finite number of steps (when $s_n = \frac{1}{l_{n+1}}$), while if s is irrational then the expansion is infinite. An infinite ceiled continued fraction expansion is written in the form $\lceil\!\lceil l_1, l_2, \ldots, l_n, \ldots \rceil\!\rceil$. A finite ceiled continued fraction is written in the form $\lceil\!\lceil l_1, \ldots, l_n \rceil\!\rceil$, while a partial ceiled continued fraction of order n of a number α is written in the form $\lceil\!\lceil l_1, \ldots, l_n - \alpha_n \rceil\!\rceil$.

Definition 7 (Ceiled convergents). *Let s be a real number such that* $0 \leq s \leq 1$ *which continued fraction is* $s = \lceil\!\lceil l_1, \ldots, l_n, \ldots \rceil\!\rceil$*. We define the sequence* $r_n(s)$ *of rational numbers* $r_n(s) = \lceil\!\lceil l_1, l_2, \ldots, l_n \rceil\!\rceil$ *by stopping the expansion and keeping the n first coefficients. These numbers are called ceiled convergents of s.*

Since $r_n(s)$ are rational numbers one can write them under the form $r_n(s) = p_n(s)/q_n(s)$, where $p_n(s)$ and $q_n(s)$ could be computed for $n \geq 2$ with the following recurrence relation :

$$p_n(s) = l_n p_{n-1}(s) - p_{n-2}(s) \quad \text{and} \quad q_n(s) = l_n q_{n-1}(s) - q_{n-2}(s),$$

with initial conditions $p_0(s) \stackrel{\triangle}{=} 0$, $p_1(s) \stackrel{\triangle}{=} 1$, $q_0(s) \stackrel{\triangle}{=} 1$ and $q_1(s) \stackrel{\triangle}{=} l_1$.

The sequence of ceiled convergents of s converge to s optimally from below: $r_n(s)$ is the rational number with the smallest denominator in the interval $]r_{n-1}(s), s]$. Furthermore, the sequence of ceiled convergents define a sequence of Farey's intervals.

Note that the even classical convergents form a subsequence of the ceiled convergents.

3 Factorization of mechanical words with ceiled continued fraction

While the factorizations of mechanical words with classical continued fraction is well known and has been developed in numerous works (see [4]), the relation between the ceiled continued fractions and mechanicals word presented here is new and has several advantages as shown later. From now on, we focus on mechanical words with intercept 0.

Let $\overline{M}$ be the set of all upper mechanical words and $\overline{M_k}$ be the set of all mechanical words of slope α such that $(k+1)^{-1} \leq \alpha < k^{-1}$. From Lemma 3, all the words of $\overline{M_k}$ have either $k-1$ or k consecutive zeros between any two ones.

The morphism $\overline{\varphi}_k : \overline{M}_k \to A^* \cup A^{\mathbb{N}}$, replaces the maximal sub-words which begin by 1 and then only contain 0's according to the following rule :

$$\overline{\varphi}_k : \begin{array}{lcl} 1\overbrace{0\ldots0}^{k-1} & \mapsto & 1 \\ 1\overbrace{0\ldots0}^{k} & \mapsto & 0 \end{array} ,$$

if $\alpha = (k+1)^{-1}$ then the word $\overline{\varphi}_k(\overline{m}_\alpha)$ is reduced to zero. For example, $\overline{m}_{3/7} = 1010100 \in \overline{M}_2$ thus $\overline{\varphi}_2(\overline{m}_{3/7}) = 110$.

In order to extend the domain of $\overline{\varphi}_k$ to the set $\overline{M}$, consider the morphism $\overline{\Phi}$:

$$\overline{\Phi}: \begin{array}{ccc} \overline{M} & \to & A^* \cup A^{\mathbb{N}} \\ \overline{m}_\alpha & \mapsto & \overline{\varphi}_{\lfloor \alpha^{-1} \rfloor}(\overline{m}_\alpha) \end{array} .$$

Note that the morphisms defined above are different of the classical morphisms on words. Indeed it replaces groups of letters by one letter while usual morphisms replace single letters by groups of letters.

Lemma 8. *Let $\overline{m}_\alpha$ be the upper mechanical word of slope α, then $\overline{\Phi}(\overline{m}_\alpha)$ is the upper mechanical word of slope $\lceil \alpha^{-1} \rceil - \alpha^{-1}$.*

Proof. If α^{-1} is an integer number, then $\overline{\Phi}(\overline{m}_\alpha)$ is reduced to the word "0", and the result follows since $\lceil \alpha^{-1} \rceil - \alpha^{-1} = 0$.

If α^{-1} is not an integer number, since $\overline{m}_\alpha(n)$ is defined by Formula (1) we know the sequence of the indexes of letter 1. Let i_k be an integer number such that i_k is the index of the kth occurrence of the letter one in $\overline{m}_\alpha$ (assuming the first one is the 0th occurrence). This implies that $\overline{m}_\alpha(0)\ldots\overline{m}_\alpha(i_k)$ contains $k+1$ letters one and $\overline{m}_\alpha(0)\ldots\overline{m}_\alpha(i_k-1)$ contains k letters one. This means

$$\lceil \alpha(i_k+1) \rceil = k+1,\ \lceil \alpha i_k \rceil = k.$$

These equalities imply $i_k = \lfloor k\alpha^{-1} \rfloor$.

Now, define the morphism $\overline{\psi}_k$, which replaces maximal sub-words by a letter similarly as $\overline{\varphi}$. This morphism works as follows $\overline{\psi}_k : \overline{M_k} \to A^* \cup A^{\mathbb{N}}$.

$$\overline{\psi}_k : \begin{array}{ll} 1\overbrace{0\ldots0}^{k-1} & \mapsto 0 \\ 1\overbrace{0\ldots0}^{k} & \mapsto 1\,. \end{array}$$

Let $l = \lfloor \alpha^{-1} \rfloor$ and let $w = \overline{\psi}_l(\overline{m}_\alpha)$. The word w describes the sequence of the numbers of consecutive 0's between two consecutive 1's in $\overline{m}_\alpha$. The letter 1 in w corresponds to the largest number of consecutive 0's in $\overline{m}_\alpha$ while the letter 0 in w corresponds to the smallest number of consecutive 0's in $\overline{m}_\alpha$. Hence, the sequence $(w(k))_{k\geq 0}$ of letters of w can be computed by $w(k) = i_{k+1} - i_k - l$. Then $w(k) = \lfloor (k+1)\alpha^{-1} - l(k+1) \rfloor - \lfloor k\alpha^{-1} - kl \rfloor$, and therefore w is the lower mechanical word of slope $\alpha^{-1} - l$

Define now the morphism γ such that

$$\gamma: \begin{array}{ccc} 1 & \mapsto & 0 \\ 0 & \mapsto & 1\,. \end{array}$$

From [15], it is known that the morphism γ transforms a lower mechanical word of slope α into an upper mechanical word of slope $1-\alpha$ and conversely. Let $w' = \gamma(w)$, w' is the upper mechanical word of slope $1-(\alpha^{-1}-l)$.

It can be checked that $\overline{\Phi} = \gamma \circ \overline{\psi}$, hence $\overline{\Phi}(\overline{m}_\alpha)$ is the upper mechanical word of slope $1-(\alpha^{-1} - \lfloor \alpha^{-1} \rfloor) = \lceil \alpha^{-1} \rceil - \alpha^{-1}$. □

The following corollary shows the connection between $\Phi(\overline{m}_\alpha)$ and the ceiled continued fraction of the slope of $\overline{m}_\alpha$.

Corollary 9. *Let α, $0<\alpha<1$ be any real number which satisfies $\alpha = [\![l_1,\ldots,l_{n-1},l_n-\alpha_n]\!]$, with $n \geq 1$. If $m = \Phi(\overline{m}_\alpha)$ then the slope of m satisfies $\mathbf{s}(m) = [\![l_2,l_3,\ldots,l_{n-1},l_n-\alpha_n]\!]$.*

Proof. From (4), it comes $\alpha = (l_1-\alpha_1)^{-1}$ and $\alpha_1 = [\![l_2,l_3,\ldots,l_n-\alpha_n]\!]$. With Lemma 8, we get $\mathbf{s}(m) = \lceil \alpha^{-1} \rceil - \alpha^{-1} = l_1 - \alpha^{-1} = \alpha_1$. Therefore $\mathbf{s}(m) = [\![l_2,l_3,\ldots,l_{n-1},l_n-\alpha_n]\!]$. □

The previous corollary can be re-expressed as, if $\alpha = [\![l_1-\alpha_1]\!]$, then $\overline{\Phi}(\overline{m}_\alpha) = \overline{m}_{\alpha_1}$. These results allow one to apply a recurrence relation on the number of terms in a continued fraction expansion.

Theorem 10 (($\overline{x}$,$\overline{y}$)-factor decomposition). *Let α, $0<\alpha<1$ be any given real number with $\alpha=\lceil\!\lceil l_1,l_2,\dots\rceil\!\rceil$ and consider the two sequences, $\{\overline{x}_i(\alpha)\}_{i\geq 0}$ and $\{\overline{y}_i(\alpha)\}_{i\geq 0}$, computed by :*

$$\begin{aligned} \overline{x}_0(\alpha)&=1, & \overline{x}_i(\alpha)&=\overline{x}_{i-1}(\alpha)\big(\overline{y}_{i-1}(\alpha)\big)^{l_i-2}, & &\textit{for } i\geq 1,\\ \overline{y}_0(\alpha)&=0, & \overline{y}_i(\alpha)&=\overline{x}_{i-1}(\alpha)\big(\overline{y}_{i-1}(\alpha)\big)^{l_i-1}, & &\textit{for } i\geq 1. \end{aligned} \tag{5}$$

Then the upper mechanical word $\overline{m}_\alpha$ can be factorized only using the two factors $\overline{x}_i(\alpha)$ and $\overline{y}_i(\alpha)$. These two sequences are called ($\overline{x}$-$\overline{y}$)-factor decomposition sequences associated with the mechanical expansion of α.

Proof. The finite sequences $\{\overline{x}_i(\alpha)\}_{0\leq i\leq n}$, $\{\overline{y}_i(\alpha)\}_{0\leq i\leq n}$ are associated with the partial mechanical continued fraction expansion of α and this will prove the result by induction.

Step 1. $\alpha=\lceil\!\lceil l_1-\alpha_1\rceil\!\rceil$. Since, from Lemma 3, the number of consecutive 0 between to 1 can take only two values, $\overline{m}_\alpha$ can be factorized with $\overline{x}_1(\alpha)$ and $\overline{y}_1(\alpha)$.

Step n. We have $\alpha=\lceil\!\lceil l_1,l_2,\dots,l_n-\alpha_n\rceil\!\rceil$. Let $w=\overline{\Phi}(\overline{m}_\alpha)$, by Corollary 9, $\mathbf{s}(w)=\lceil\!\lceil l_2,l_3,\dots,l_{n-1},l_n-\alpha_n\rceil\!\rceil$. Let

$$\{\overline{x}_i(\mathbf{s}(w))\}_{0\leq i\leq n-1} \quad \text{and} \quad \{\overline{y}_i(\mathbf{s}(w))\}_{0\leq i\leq n-1}\,,$$

be the two ($\overline{x}$-$\overline{y}$)-factor decomposition sequences associated with the partial mechanical expansion of $\mathbf{s}(w)$. By induction hypothesis, w can be factorized only using $\overline{x}_i(\mathbf{s}(w))$ and $\overline{y}_i(\mathbf{s}(w))$. Introduce now the sequences $\overline{x}'_i$ and $\overline{y}'_i$ such that

$$\overline{x}'_0=1\,,\ \overline{y}'_0=0\,,\ \overline{\varphi}_{l_1-1}(\overline{x}'_i)=\overline{x}_{i-1}(\mathbf{s}(w))\,,\ \overline{\varphi}_{l_1-1}(\overline{y}'_i)=\overline{y}_{i-1}(\mathbf{s}(w))\,,\ \forall i\geq 1\,.$$

But $w=\overline{\Phi}(\overline{m}_\alpha)$ can be factorized only using $\overline{\varphi}_{l_1-1}(\overline{x}'_i)$ and $\overline{\varphi}_{l_1-1}(\overline{y}'_i)$, this $\forall i\geq 0$. Therefore $\overline{m}_\alpha$ can be factorized only using $\overline{x}'_i$ and $\overline{y}'_i$, since $\overline{\Phi}(\overline{m}_\alpha)=\overline{\varphi}_{l_1-1}(\overline{m}_\alpha)$. Note now that the sequences $\overline{x}'_i$, $\overline{y}'_i$ are the ($\overline{x}$-$\overline{y}$)-factor decomposition sequences associated with $\lceil\!\lceil l_1,l_2,\dots,l_n\rceil\!\rceil$. The use of Proposition 1 ends the proof. $\square$

Let Δ' and $\Gamma(u,v)$ be two morphisms on words such that $\Delta'(u,v)=(uv,v)$ and $\Gamma(u,v)=(u,uv)$. Then, the ($\overline{x}$,$\overline{y}$) factors verify for all $n\geq 0$, $(\Delta')^{l_n}(\overline{x}_{n-1}(\alpha),\overline{y}_{n-1}(\alpha))=(\overline{x}_n(\alpha),\overline{y}_{n-1}(\alpha))$ and $\Gamma(\overline{x}_n(\alpha),\overline{y}_{n-1}(\alpha))=(\overline{x}_n(\alpha),\overline{y}_n(\alpha))$. This implies that the factor $\overline{x}_n(\alpha)$ and the factor $\overline{y}_n(\alpha)$ form a *Christoffel pair*, since the family of *Christoffel pairs* is the smallest set of words containing $(0,1)$ and closed under Δ' and Γ, (see [15]).

When α is a rational number the factorization is stopped and the upper mechanical word can be factorized under a special form

Lemma 11. *Let α, $0<\alpha<1$ be any given rational number with $\alpha=\lceil\!\lceil l_1,\dots,l_n\rceil\!\rceil$. Then $\overline{y}_n(\alpha)=\overline{m}_{r_n(\alpha)}=\overline{m}_\alpha$.*

Proof. The proof holds by induction on the number of terms in the finite mechanical expansion.

Step 1. Considering that $\alpha=\lceil\!\lceil l_1\rceil\!\rceil$ then α is a rational number with $\alpha^{-1}=l_1$. Hence $\overline{m}_\alpha=10\dots0$, where the number of consecutive 0 is $\alpha^{-1}-1=l_1-1$. Thus $\overline{m}_\alpha=y_1(\alpha)$ and the induction hypothesis is proved.

Step n. We have $\alpha=\lceil\!\lceil l_1,l_2,\dots,l_{n-1},l_n\rceil\!\rceil$, therefore the word w defined by $w=\overline{\Phi}(\overline{m}_\alpha)$ has a slope $\mathbf{s}(w)=\lceil\!\lceil l_2,\dots,l_n\rceil\!\rceil$. Using the induction hypothesis we obtain $w=\overline{y}_{n-1}(\mathbf{s}(w))$ and $w=\overline{m}_{r_{n-1}(\mathbf{s}(m)}$. Since $\overline{y}_{n-1}(\mathbf{s}(w))=\overline{\varphi}_{l_1-1}(\overline{y}_n(\alpha))$, then $\overline{m}_\alpha=\overline{y}_n(\alpha)$. $\square$

Example 12. *Let $\alpha=\frac{11}{29}$. With the use of Equations (4), the ceiled continued fraction of α is $\lceil\!\lceil 3,3,4\rceil\!\rceil$, and $\overline{m}_{11/29}=10100101001001010010010100100$. Then $\Phi(\overline{m}_{11/29})=10100100100$, $\Phi^{(2)}(\overline{m}_{11/29})=1000$, and $\Phi^{(3)}(\overline{m}_{11/29})=0$.*

The construction of the ($\overline{x}$,$\overline{y}$) factors gives

$$\overline{m}_{11/29}=\overbrace{\underbrace{\overbrace{10}^{\overline{x}_1}\,\overbrace{100}^{\overline{y}_1}}_{\overline{x}_2}\,\underbrace{\overbrace{10}^{\overline{x}_1}\,\overbrace{100}^{\overline{y}_1}\,\overbrace{100}^{\overline{y}_1}}_{\overline{y}_2}\,\underbrace{\overbrace{10}^{\overline{x}_1}\,\overbrace{100}^{\overline{y}_1}\,\overbrace{100}^{\overline{y}_1}}_{\overline{y}_2}\,\underbrace{\overbrace{10}^{\overline{x}_1}\,\overbrace{100}^{\overline{y}_1}\,\overbrace{100}^{\overline{y}_1}}_{\overline{y}_2}}^{\overline{y}_3}\,.$$

Remark 13. *Since* $\Phi^{(n)}(\overline{m}_\alpha) = \varphi_{l_n-1} \circ \ldots \circ \varphi_{l_2-1} \circ \varphi_{l_1-1}(\overline{m}_\alpha)$, *then* $\overline{x}_n(\alpha)$ *and* $\overline{y}_n(\alpha)$ *are the only factors of* $\overline{m}_\alpha$ *verifying*

$$\Phi^{(n)}(\overline{x}_n(\alpha)) = 1 \quad \text{and} \quad \Phi^{(n)}(\overline{y}_n(\alpha)) = 0.$$

From this, it can be deduced that the slope of $\Phi^{(n)}(\overline{m}_\alpha)$ *represents the frequency of* $\overline{x}_n(\alpha)$ *in* $\overline{m}_\alpha$.

Also note that $\overline{x}_i(\alpha)$ *is a factor with a larger slope than* α *while* $\overline{y}_i(\alpha)$ *is a factor with a smaller slope than* α.

Corollary 14. *Let* $\overline{y}_n(\alpha)$ *be the sequence of factors computed by Formula* (5)*, then the sequences of* $y_n(\alpha)$ *is a sequence of suffixes of* $\overline{m}_\alpha$ *which, furthermore, for any* $\alpha \in [0,1]$ *satisfies* $\lim_{n\to\infty} \overline{y}_n(\alpha) = \overline{m}_\alpha$.

Proof. The proof follows of Lemma 11 and Remark 13. □

Finally note that all the results presented in this section also hold for lower mechanical words by inverting the order of the factors $\overline{x}_n(\alpha)$ and $\overline{y}_n(\alpha)$ at each step.

3.1 Ceiled factorization of the characteristic words

The ($\overline{x}$,$\overline{y}$)-factor decomposition allows to have a finite sequence of prefixes of the characteristic word.

Since $0c_\alpha = \underline{m}_\alpha 0$ when α is rational and since $0c_\alpha = \underline{m}_\alpha$ otherwise, we deal with lower mechanical words, in order to keep the results more general. Indeed, we handle only with prefixes of c_α. Thus, the factors $\overline{x}$ and $\overline{y}$ are now these ones associated with the factorization of the lower mechanical word. For notation simplicity, the indication of the slope is skipped.

Corollary 15. *Let* z_n *be a sequence of factors defined by* $z_n = \overline{x}_1\overline{x}_2 \ldots \overline{x}_n$. *Then, the sequence* z_n *is a sequence of prefixes of* c_α, *which for any irrational* α *satisfies* $\lim_{n\to\infty} z_n = c_\alpha$.

Proof. Let m_n be the word such that $0m_n = \overline{y}_n$. We claim that $m_n = \overline{x}_1\overline{x}_2 \ldots \overline{x}_n$. The proof of the claim is made by induction. At step 1 by (5) we have $m_1 = \overline{x}_1$. At step n, it comes

$$\overline{y}_n = \overline{y}_{n-1}(\overline{y}_{n-1})^{l_n-2}\overline{x}_{n-1} = 0m_{n-1}\overline{x}_n = 0\overline{x}_1\overline{x}_2 \ldots \overline{x}_n.$$

Noticing that the $\overline{y}_n$ form a sequence of prefixes of $\underline{m}_\alpha$ and since $0c_\alpha = \underline{m}_\alpha$ ends the proof. □

Note that, $\overline{\Phi}(0z_n) = 0$ and furthermore if v_n is the word such that $\underline{m}_\alpha = 0z_nv_n$, then $\overline{\Phi}(v_n) = c_{\alpha_n}$.

Example 16 (Fibonacci word). *We present now the factorization of the characteristic word of slope* $\alpha = (1+\sqrt{5})/2$. *The ceiled expansion of* α *is* $⟦2,3,3\ldots⟧$. *Henceforth,*

$$c_\alpha = \underbrace{\underbrace{\overbrace{01}^{\overline{x}_1}\,\overbrace{00101}^{\overline{x}_2}}_{z_2}\overbrace{001001010010}^{\overline{x}_3}}_{z_3}\ldots$$

4 Factorization with Classical Continued Fraction

This section can be seen as the dual of Section 3, in the sense that it connects the factorization of mechanical words with a morphism which, in turn is correlated to the classical continued fraction. Moreover, this morphism leads to factors with similar properties as those presented in Remark 13. For simplicity, the results are presented on lower mechanical words only, but could also be obtained for upper mechanical words.

$\underline{M}$ is the set of all lower mechanical words and $\underline{M}_k$ the set of all lower mechanical words whose slope α verifies $(k+1)^{-1} \le \alpha < k^{-1}$. The morphism $\underline{\Phi}$ from $\underline{M} \mapsto \underline{M}$ is defined by $\underline{\Phi} = \gamma \circ \overline{\Phi} \circ \gamma$.

Lemma 17. *Let* $\alpha \in [0,1]$ *be such that* $\alpha = [0, a_1 + \alpha_1]$. *The morphism* $\underline{\Phi}$ *changes the lower mechanical word of slope* $1-\alpha$ *into the lower mechanical word of slope* α_1.

Proof. Let m be the word such that $m=\gamma(\underline{m}_{1-\alpha})$, it is known that $m=\overline{m}_\alpha$. Let $m'=\overline{\Phi}(m)$, then $m'=\overline{m}_{\alpha'_1}$ where α'_1 is the remainder of the ceiled continued fraction at the first order of the expansion, *i.e.* $\alpha=\lceil\!\lceil l_1-\alpha'_1\rceil\!\rceil$. Let $m''=\gamma(m')$ then $m''=\underline{m}_{1-\alpha'_1}$. By rewriting $1-\alpha'_1$ in function of α, and using (4), $\alpha'_1=\lceil\frac{1}{\alpha}\rceil-\frac{1}{\alpha}$, one gets $1-\alpha'_1=\frac{1}{\alpha}-\left(\lceil\frac{1}{\alpha}\rceil-1\right)$, yielding $1-\alpha'_1=\frac{1}{\alpha}-\lfloor\frac{1}{\alpha}\rfloor$. Thus $1-\alpha'_1=\alpha_1$. □

The morphism $\underline{\Phi}$ is an the extension of the morphism $\underline{\phi}^k$ to $\underline{M}$ by $\underline{\Phi}(\underline{m}_\alpha)=\underline{\phi}^{\lceil(1-\alpha)^{-1}\rceil}(\overline{m}_\alpha)$. Since the words $\underline{m}_\alpha$ and $\overline{m}_\alpha$ are mechanical, then the number of consecutive ones in $\underline{m}_\alpha$ and $\overline{m}_\alpha$ can only take two values. More precisely, the maximum number of consecutive ones is equal to $\lceil(1-\alpha)^{-1}\rceil-1$, while the minimum number of consecutive ones is equal to $\lfloor(1-\alpha)^{-1}\rfloor-1$. This can be deduced from Lemma 3 using the fact that the number of consecutive ones of a word m is equal to the number of consecutive zeros in $\gamma(m)$. Hence, the morphism $\underline{\Phi}$ deals with words with consecutive one, while $\overline{\Phi}$ deals with words with consecutive 0.

Consider the following example: $\underline{m}_{3/11}=00010001001$, this means that $\underline{\Phi}(\underline{m}_{3/11})=\underline{\phi}^1(\underline{m}_{3/11})=00100101=\underline{m}_{3/8}$. Also consider $\underline{m}_{5/7}=0110111$, then $\underline{\Phi}(\underline{m}_{5/7})=\underline{\phi}^3(\underline{m}_{5/7})=01$.

We can deduce then

Theorem 18 (($\underline{x}-\underline{y}$) factorization of lower mechanical words). *Let $\alpha\in[0,1]$ such that its classical continued fraction is $[a_1,\ldots,a_n,\ldots]$. We define now the two sequences of factors $\{\underline{x}_n(\alpha)\}_{n\in\mathbb{N}}$ and $\{\underline{y}_n(\alpha)\}_{n\in\mathbb{N}}$ computed by $\underline{x}_0=0$, $\underline{y}_0=1$,*

$$\underline{x}_n=\underline{y}_{n-1}^{a_n-1}\underline{x}_{n-1},\quad \underline{y}_n=\underline{y}_{n-1}^{a_n}\underline{x}_{n-1},\qquad \textit{if } n \textit{ is odd,} \tag{6}$$

$$\underline{x}_n=\underline{y}_{n-1}\underline{x}_{n-1}^{a_n},\quad \underline{y}_n=\underline{y}_{n-1}\underline{x}_{n-1}^{a_n-1},\qquad \textit{if } n \textit{ is even} \tag{7}$$

Then the word $\underline{m}_\alpha$ can be factorized only by the factors $\underline{x}_n(\alpha)$ and $\underline{y}_n(\alpha)$.

Proof. Let α be a real number such that $\alpha=[a_1+\alpha_1]=[a_1,a_2+\alpha_2]=\ldots=[a_1,a_2,\ldots,a_n+\alpha_n]$. By Lemma 8, we know that the application of the morphism $\overline{\Phi}$ transforms $\underline{m}_\alpha$ in a lower mechanical word which slope is $1-\alpha_1$. From Lemma 17, the morphism $\underline{\Phi}$ transforms $\underline{m}_{1-\alpha_1}$ in a lower mechanical word of slope α_2 where $\alpha_1=[a_2+\alpha_2]$. Hence we have $\underline{\Phi}\circ\overline{\Phi}(\underline{m}_\alpha)=\underline{m}_{\alpha_2}$. Applied to any even order of the expansion in continued fraction this composition morphisms gives a lower mechanical word which slope is an even remainder of the continued fraction.

Lemma 8 and Lemma 17 as well as the definition of $\underline{\Phi}$ show that $\underline{x}_i(\alpha)$ and $\underline{y}_i(\alpha)$ are sequences built such that the factors $\underline{x}_i(\alpha)$ are represented by the letter 1 and such that the factors $\underline{y}_i(\alpha)$ are represented by the letter 0 in any composition of the morphism $\overline{\Phi}^i\circ(\underline{\Phi}\circ\overline{\Phi})^{(j)}(\underline{m}_\alpha)$ with i equal to 0 or 1 and $j\in\mathbb{N}$, assuming that $\overline{\Phi}^0$ is the identity.

The use of a similar induction as for Theorem 10 ends the proof. □

Remark 19. *One other interest of the two factorizations presented here is related to computer graphics especially the display of straight lines. Indeed, the maximal and the minimal consecutive either* 0 *or* 1 *in words is called the span and the crucial points in the field concerns the manner to distribute the maximal and the minimal spans in the word. We think that this can be done using the factorization here while it was not efficient with former factorizations of Sturmian words (see [3]).*

Example 20 (Factorization of $\underline{m}_{11/29}$). *In this example we will consider the lower mechanical word of slope* $11/29=[2,1,1,1,3]$ *:* $\underline{m}_{11/29}=00100101001001010010010100101$. *Then, we have* $\overline{\Phi}(\underline{m}_{8/19})=\underline{m}_{4/11}=00100100101$, $\underline{\Phi}(\underline{m}_{4/11})=\underline{m}_{4/7}=0101011$, $\overline{\Phi}(\underline{m}_{4/7})=\underline{m}_{1/4}=0001$, $\underline{\Phi}(\underline{m}_{1/4})=\underline{m}_{1/3}=001$, $\underline{\Phi}(\underline{m}_{1/3})=$

$\underline{m}_1 = 1$. *The ($\underline{x}$-$\underline{y}$) factorization of the lower mechanical word is*

$$\underline{m}_{11/29} = \underbrace{\overbrace{\underbrace{\overbrace{\underbrace{001}_{\underline{y}_1}}^{\underline{y}_2}\overbrace{\underbrace{001}_{\underline{y}_1}\underbrace{01}_{\underline{x}_1}}^{\underline{x}_2}}_{\underline{y}_3}}^{\underline{y}_4}\overbrace{\underbrace{\overbrace{\underbrace{001}_{\underline{y}_1}}^{\underline{y}_2}\overbrace{\underbrace{001}_{\underline{y}_1}\underbrace{01}_{\underline{x}_1}}^{\underline{x}_2}}_{\underline{y}_3}}^{\underline{y}_4}\overbrace{\underbrace{\overbrace{\underbrace{001}_{\underline{y}_1}}^{\underline{y}_2}\overbrace{\underbrace{001}_{\underline{y}_1}\underbrace{01}_{\underline{x}_1}}^{\underline{x}_2}}_{\underline{y}_3}\underbrace{\overbrace{\underbrace{001}_{\underline{y}_1}\underbrace{01}_{\underline{x}_1}}^{\underline{x}_2}}_{\underline{x}_3}}^{\underline{x}_4}}_{\underline{x}_5}$$

Note the factorization requires, here, 5 *steps while the factorization of section 3 requires only* 3 *steps.*

The factors are at each order of the expansion built differently because of the ordering of convergents of the classical continued fraction. The even convergents of a number s form an an increasing sequence and the odd convergents form a decreasing sequence.

Remark 21. *Corollary 14 does not have a counterpart. Indeed, if N denotes the last order of a classical expansion of a rational number α, then if N is even, $\underline{\Phi}^N = 0$ and if N is odd, $\underline{\Phi}^N = 1$. This means that $\underline{m}_\alpha = \underline{y}_N$ for an even N and $\underline{m}_\alpha = \underline{x}_N$ for an odd N.*

4.1 Connection with classical factorisations

There exists a connection between this factorization and that presented in [15] for mechanical words.

Theorem 22 ([15]). *Let $\alpha = [a_1 - 1, a_2, \ldots]$ then the sequence of words u_n built by $u_0 = 0$, $u_{-1} = 1$, $u_1 = (u_0)^{a_1 - 1} u_1$,*

$$u_n = (u_{n-1})^{a_n} u_{n-2} \ \textit{if } n \textit{ is odd}, \quad u_n = u_{n_2} (u_{n-1})^{a_n} \ \textit{if } n \textit{ is even},$$

is a sequence of factors of $\underline{m}_\alpha$ such that $\lim_{n\to\infty} u_n = \underline{m}_\alpha$.

Lemma 23. *Let u_n be the sequence of factor defined above, then $u_n = \underline{x}_n$ when n is odd and $u_n = \underline{y}_n$ when n is even.*

Proof. We have by (6) $u_0 = \underline{y}_0$ and $u_1 = \underline{x}_1$. Assume that the result is true still for $n-1$ and that n is even. We have

$$u_n = u_{n_2}(u_{n-1})^{a_n} = \underline{y}_{n-2}(\underline{x}_{n-1})^{a_n} = \underline{y}_{n-2}\underline{x}_{n-1}(\underline{x}_{n-1})^{a_n - 1} = \underline{y}_{n-1}(\underline{x}_{n-1})^{a_n - 1},$$

hence $u_n = \underline{y}_n$. The odd case is similar. □

An new (simple) proof of Theorem 22 follows immediately from the Lemma 23 and Theorem 18. It can be also deduced from these two results that for any $n \geq 1$ we have $u_n = \underline{m}_{\rho_n}$, but in contrary to of the factorization presented section 3 this can not be associated with sequences of either prefixes or suffixes.

4.2 Comparison of speeds of convergence

As seen in Example 20, the number of steps required to approximate a number differs according to the continued fraction expansion chosen. We give here a brief comparison of these numbers of steps for each of the continued fraction.

Definition 24 (Intermediate convergent). *Let α be a given real number in $[0,1]$. Let $\rho_{2n} = [a_1, \ldots, a_{2n}]$ be the sequence of the even classical convergents of α. We define as intermediate convergents the rational numbers given by $[a_1, \ldots, a_{2n-1}, j]$, with $j \in \{1, \ldots, a_{2n}\}$.*

The sequence of intermediate convergents and even convergents of α is the sequence of ceiled convergents r_n of α. Indeed, from [12], the nth ceiled convergent of α is the rational number with smallest denominator in $[r_{n-1}(\alpha),\alpha]$. On the other hand, using Farey's properties of classical convergents, at order n the jth intermediate convergents is the rational number of smallest denominator in the interval $[[a_1,\ldots,a_{2n-1},j],\rho_{2n-1}]]$ that is the rational number of smallest denominator in $[[a_1,\ldots,a_{2n-1},j],\alpha$. Noticing that, by definition, $a_1+1=l_1$ concludes the claim.

Since the number of intermediate convergents is the number of ceiled convergents between two classical even convergents, then the differences of numbers of steps between ceiled continued fraction and classical continued fraction can be computed and depend only on the value of the even coefficients of the classical continued fraction.

Lemma 25. *Let α be a number in $[0,1]$ such that $\alpha=[a_1,a_2,\ldots,a_n+\alpha_n]$, where α_n can be null. If,*

$$\sum_{i=1}^{\lfloor n/2\rfloor}\mathbf{1}_{\{a_{2i}=1\}}\geq\sum_{i=1}^{\lfloor n/2\rfloor}(a_{2i}-1),$$

then it is most efficient to use a ceiled continued fraction expansion.

Proof. We focus on a given even order : $2i$, since the rest of the proof follows. Let assume that $\rho_{2(i-1)}=r_{2(i-1)}$. If $a_{2i}=1$, then the number of intermediate convergents is 0 and $\rho_{2i}=r_{2i-1}$. If $a_{2i}=2$, then the number of intermediate convergents is 1 and $\rho_{2i}=r_{2i}$. If $a_{2i}=j$, with $j\geq 3$, then the number of intermediate convergents is $j-1$ and $\rho_{2i}=r_{2i+j-1}$. $\square$

However, the most efficient expansion cannot be guessed *a priori* and one has to compute the classical expansion, in order to know which expansion is the most efficient.

Remark 26. *Using the previous lemma it could be noticed that the factorization of the Fibonacci word $c_{(1+\sqrt{5})/2}$ presented in Example 16 is twice faster than the classical one, since $(1+\sqrt{5})/2=[1,1,\ldots]$.*

5 Applications

We present now some algorithmic applications of the morphism $\overline{\Phi}$. Indeed, this morphism is different from the classical morphisms on words because it replaces groups of letters by one letter while usual morphisms replace single letters by groups of letters. This remark is the starting point of several applications of this morphism to construction problems. All algorithms provided in the following have optimal (linear) asymptotic complexity.

5.1 A linear test for periodic mechanical words

The morphism $\overline{\Phi}$ can be used to check if a finite word w is mechanical in linear time ($O(|w|)$). The outline of this algorithm is given in Figure 5.1.

Correctness and linearity of the algorithm The correctness of the algorithm can be deduced from Lemma 8 and Corollary 9.

The linearity comes from the fact that the word obtained in step 4, is at least twice shorter as previously. Since, during steps 1, 2, and 3 the number of operations is linear with the length of the word treated, then the total complexity is smaller than $K|w|\sum_{n=1}^{n=\lceil|w|/2\rceil}\frac{1}{2^n}$, where K is a finite constant. Thus the total complexity is smaller than $2K|w|$ and the algorithm is linear.

Note that some improvements can be achieved on the constant K by merging operations during steps 1, 2 and 3.

1. input: w

2. If $|w| \geq 1$, scan the word w to compute its slope α.
   If $\alpha = 0$ or $\alpha = 1$, then output(*"the word is mechanical"*) and stop.

3. If $\alpha > 1/2$, then $w := \gamma(w)$ and $\alpha := 1 - \alpha$.

4. Apply a circular shift on w so that it starts with "1".
   Try to apply morphism $\overline{\varphi}_{\lfloor \alpha^{-1} \rfloor}$ on the shifted w.
   If this fails, then output(*"the word is not mechanical"*) and stop, else $w := \overline{\varphi}_{\lfloor \alpha^{-1} \rfloor}(w)$ and go to step 2.

Figure 1: Algorithm to test if a finite word is mechanical

5.2 Cellular lines

Cellular lines are finite combinatoric objects used in discrete geometry for the computations of discrete straight lines [9]. Actually, they first appeared in operations research for *Ant* algorithms [5]. A characterization in terms of words has been made in [10] .

Up to our knowledge, no fast constructions of cellular lines has been proposed yet. However, [10] provides a method to verify if a given word is a cellular line but this does not give any polynomial algorithm to compute cellular lines. This is done here using the factorization procedure defined in the previous sections.

Let $T_{\ell,r}$ be the set of all finite words made of two consecutive integers, whose length is ℓ and whose sum is r. In a word in $T_{\ell,r}$ the least frequent symbol is called the *minority* symbol while the other symbol is called the *majority* symbol. If both symbols appear the same number of times, the minority symbol is chosen arbitrarily.

Definition 27 (fully partitioned words). *A finite word is said fully partitioned if the minority symbol partitions the word into $n+1$ non empty sub-words, of majority symbols (where n is the number of occurrences of the minority symbol) or, if both symbols appear the same number of times, into n non empty sub-words.*

Let ϕ be the transformation on fully partitioned words into finite words, which replaces each maximum sub-words of consecutive majority symbols by the length of these sub-words and which delete all the minority symbols. Furthermore, the application of ϕ on the word reduced to the letter 1 gives the word 1. Let w be a finite fully partitioned word. If $\phi^j(w)$ can be constructed, it is called the *derivated word of order j* of w ([10]).

Definition 28 ([10]). *A finite word $c_{\ell,r} \in T_{\ell,r}$ is a cellular line with parameters (ℓ, r) if its derivated words exist for all orders $j \geq 0$.*

Example 29 (Cellular line $c_{6,8}$). *It is build over the integers $1 = \lfloor \frac{8}{6} \rfloor$ and $2 = \lfloor \frac{8}{6} \rfloor + 1$ since the six letters must sum to 8. Consider the fully partitioned word $w = 121121$. $\phi(121121) = 121$ is the first order derivation. The second order derivated word is $\phi(121) = 11$, the third order derivated is 2 and all the derivated words with larger orders are all 1. Thus 121121 is a cellular line of parameters 6 and 8.*

On the other hand, $w' = 121211$ is fully partitioned but the word $\phi(w') = 112$ is not fully partitioned and 121211 is not a cellular line.

Cellular lines have unusual characteristics. For example the concatenation of the cellular line $c_{c,r}$ with itself may not be the cellular line $c_{2c,2r}$. Hence, $c_{14,32} \neq c_{7,16}\, c_{7,16}$. Also note that Cellular lines are not finite mechanical words. For example, the cellular line $c_{8,10} = 11211211$ is not mechanical. However, in the following we will show that they are suffixes of finite mechanical words.

5.2.1 Determination of cellular line with mechanical words

We present now an algorithm to construct all cellular lines, which is based on the fact that ϕ and $\overline{\Phi}$ are similar except that $\overline{\Phi}$ handles the first minority symbol.

For convenience, the alphabet of two consecutive positive integers is replaced by $\mathcal{A} = \{0,1\}$. Clearly, a word over $\mathcal{A}$ can represent many cellular lines. However, all the cellular lines with the same representative in $\mathcal{A}$, have the same behavior for ϕ. The letter $\lfloor r/\ell \rfloor$ is replaced by 0 while each letter $\lfloor r/\ell \rfloor + 1$ is replaced by 1. The word obtained by such transformation has now a slope equal to $\frac{r}{\ell} - \lfloor \frac{r}{\ell} \rfloor$ and will be called cellular lines with slope $p/q = \frac{r}{\ell} - \lfloor \frac{r}{\ell} \rfloor$ in the following.

We furthermore assume, with no loss of generality, that the slope of the cellular line is less than $1/2$ (else switch the roles of 0 and 1 in the construction below).

We consider now three functions f_1, f_2 and f_3. These functions allow one to compute the output of ϕ over the slope of the derivative word and to compute the minority and majority symbols. Let f_1 be defined by : $f_1 : \mathbb{N}^2 \setminus \{(p,2p) : p \in \mathbb{N}\} \to \mathbb{N}^2$, with $f_1(p,q) = (q-p, p+1)$. Let p and q be two integer numbers with $p > q$. Let us introduce α such that $p = \alpha q + r$, $\alpha \in \mathbb{N}, r < q$. We define α_M and α_m such that $\alpha_M = \alpha$ if $r < q-r$ and $\alpha_M = \alpha+1$ if $r \geq q-r$ and such that $\alpha_m = \alpha$ if $r \geq q-r$ and $\alpha_m = \alpha+1$ if $r < q-r$. Consider then the functions f_2 and f_3 : $f_2 : \mathbb{N}^2 \to \mathbb{N}^2$, $f_3 : \mathbb{N}^2 \to \mathbb{N}^2$ defined by $f_2(p,q) = (\min(r, q-r), q)$ and $f_3(p,q) = (\alpha_M, \alpha_m)$. When $q = 2p$ then we set $(f_2 \circ f_1)(p,2p) = (0,p)$ and $(f_3 \circ f_1)(p,2p) = (1,1)$.

The special treatment of the cases $(p,2p)$ is due to the alternative definition of cellular lines when the number of majority symbols is equal to the number of minority symbols, since $\phi(c_{p,2p}) = c_{0,p}$, (or $\phi(c_{p,2p}) = c_{p,p}$).

The following lemma precises the effects of ϕ on a fully partitionned word.

Lemma 30. *Let w be a fully partitioned word in $\mathcal{A}$ of slope $p/q < 1/2$ such that its first derivated word w' exists and is fully partitioned. After replacement in w' of the minority (resp. majority) symbol by* 1 *(resp* 0*), then the slope of w' is p'/q' where $(p',q') = (f_2 \circ f_1)(p,q)$.*

Proof. Since w is fully partitioned and since $2p < q$ then w is composed of $p+1$ sets of majority symbol : 0. The number of 0 is equal to $q-p$. Since w and w' are fully partitioned then w' is composed of two kinds of sets of letters containing either α or $\alpha+1$ letters, where α is the quotient $q-p/(p+1)$. The number of sets containing $\alpha+1$ letters is equal to the remainder of the division : r. Noticing that the $|w'| = p+1$ and that the number of 1 after replacement is $\min(r, q-p-r)$ gives the result. $\square$

Lemma 31. *If $p/q < 1/2$, then, there exists $n \leq q$ such that $(f_2 \circ f_1)^n(p,q) = (0,a)$.*

Proof. Let $(f_2 \circ f_1)(p,q) = (u,v)$. Then $u < v$ and $v < q$. This shows the lemma. $\square$

5.2.2 Algorithm

Here is an algorithm to construct a cellular line using $\overline{\Phi}$.

1. $(0,a) := (f_2 \circ f_1)^n(p,q)$.
2. $m_{n+1} := 0^a$.
3. for k from n to 1 do
   (a) $(\alpha_M, \alpha_m) := f_3 \circ f_1 \circ (f_2 \circ f_1)^{k-1}(p,q)$
   (b) • if $\alpha_M \geq \alpha_m$: $m_k = \overline{\phi}_{\alpha_M}^{-1}(m_{k+1})$,
       • if $\alpha_M < \alpha_m$: $m_k = \overline{\phi}_{\alpha_M}^{-1}(\gamma(m_{k+1}))$.
   (c) remove the first 1 in m_k, except if the slope is equal to $1/2$.

Theorem 32 (Correctness of the algorithm). *The output m_1 of the algorithm above is a cellular line with parameters (p,q). Moreover, the time complexity of this construction is $O(|m_1|)$.*

Proof. From Lemmas 30 and 31 then all m_k are fully partitioned. The decomposition of a cellular line starting from m_1 finds iteratively the words $m_2,\dots,m_n$ where 0 replaces the majority symbol and 1 the minority symbol. Since all m_k are fully partitioned and are the derivative words of m_1, therefore m_1 is the cellular line with parameters p,q.

Similar arguments as in Section 5.1 insure the linear complexity. □

Actually, the algorithm presented above can be adapted to construct *all* cellular lines. Indeed, by [10] for a given rate p/q there exists either one or two cellular lines. The number of cellular lines is two, if and only if $(f_2\circ f_1)^{n-1}(p,q)=(k,2k)$ with $k\in\mathbb{N}$. In these cases, two cellular lines with the same parameter exist, one using $m_{n-1}=\gamma\big(\overline{\phi}_{\alpha_M}^{-1}(m_n)\big)$ and one using $m_{n-1}=\overline{\phi}_{\alpha_m}^{-1}(m_n)$.

Note furthermore, that all the words m_k with $2\le k\le n+1$ are also cellular lines of parameters $(f_2\circ f_1)^{k-1}(p,q)$.

Corollary 33. *All cellular lines are suffixes of mechanical words (a fortiori a cellular line is balanced).*

Proof. Use the same construction as before without step (c) that removes the first symbol. This corresponds to the application of $\overline{\Phi}^{-1}$. Since the morphism $\overline{\Phi}$ transforms a mechanical word into a mechanical word then $\overline{\phi}^{-1}$ does the same. Thus, the output of this last construction is mechanical since the initial state is. □

The slope of the mechanical word obtained using the construction presented in Corollary 33 depends on the whole construction process and cannot be guessed *a priori*. Note that if γ is never used then one gets an upper mechanical word.

Example 34. *Construction of the cellular line with slope* $4/13$*. The consecutive uses of* f_1*,* f_2 *and* f_3 *give*

$$\begin{array}{lll} f_1(4,13)=(9,5), & f_2(9,5)=(1,5), & f_3(9,5)=(2,1),\\ f_1(1,5)=(4,2), & f_2(4,2)=(0,2), & f_3(4,2)=(2,0), \end{array}$$

We get $m_3=00$*. Since at orders* 2 *and* 3 $\alpha_M>\alpha_m$ *then* $m_2=\overline{\phi}_2^{-1}(00)$ *with the first one removed:* $m_2=00100$ *and* $m_1=\overline{\phi}_2^{-1}(00100)$*, hence the cellular line is* $m_1=0010010100100$*.*

For the cellular line of slope $4/13$*, the mechanical word obtained with the method presented in Corollary 33 gives two periods of* $\overline{m}_{3/8}$ *:* $\overline{m}_{3/8}\cdot\overline{m}_{3/8}=1010010010100100$*. Finally,* 10010010100100 *is also a mechanical word (with slope* $5/14$*) with the cellular line of slope* $4/13$ *as a suffix. Moreover it is shorter than* $\overline{m}_{3/8}\cdot\overline{m}_{3/8}$ *and has a better slope:* $|5/14-4/13|<|3/8-4/13|$*. Computing the shortest mechanical word for which a given cellular line is a suffix (or a prefix) seems to be a more difficult problem.*

6 Conclusion

This paper presents a morphism on Christoffel words. This morphism allows one to factorize upper and lower mechanical words of intercept 0 and to extend a previous result on the factorization of characteristic words. It also allows us to characterize and to build combinatorial object called Cellular lines. The use of these objects in optimization problems with Ant algorithms, their characterizations by mechanical words as well as the role played by mechanical words in operation research problems [7] call for further investigations for possible contributions in the optimization field, of this morphism and of Christoffel words.

References

[1] J.-P. Allouche. Algebraic and analytic randomness. In M. Planat, editor, *Noise, Oscillators and algebraic randomness*, Lecture Notes in Physics, pages 345–356. Springer-Verlag, 2000.

[2] J.-P. Allouche, J. L. Davison, M. Queffélec, and L. Q. Zamboni. Transcendence of sturmian or morphic continued fractions. *J. Number Theory*, 91:39–66, 2001.

[3] V. Boyer and J.L. Bourdin. Fast lines: a span by span method. In *Eurographics 99*, 1999.

[4] T.C. Brown. Descriptions of the characteristic sequence of an irrational. *Canadian Mathematics Bulletin*, 36:15–21, 1993.

[5] M. Dorigo, V. Maniezzo, and A. Colorni. The ant system : Optimization by a colony of cooperating agents. *IEEE Transactions on System, Man and Cybernetics*, 26(1):1–13, 1996.

[6] P. Flajolet. Combinatorial aspects of continued fractions. *Discrete Mathematics*, 32:125–161, 1980.

[7] B. Gaujal and E. Hyon. Optimal routing policies in two deterministic queues. *Réseaux et systèmes répartis - Calculateurs Parallèles*, 13(6):601–633, 2001.

[8] B. Gaujal and E. Hyon. Optimal routing policies in deterministic queues in tandem. In *WODES*, pages 251–257. IEEE, 2002.

[9] P. Geer, H.W. McLaughlin, and K. Unsworth. Discrete lines and ant algorithm. Technical report, Lincoln University, 2001.

[10] P. Geer, H.W. McLaughlin, and K. Unsworth. Cellular line : An introduction. In *2nd Conference of Discrete Models for Complex Systems*, pages 167–178, 2003.

[11] I. Goulden and D. Jackson. *Combinatorial Enumeration*. Wiley, 1983.

[12] E. Hyon. *Contrôle d'admission en boucle ouverte dans les réseaux*. PhD thesis, INPL, 2002. in French.

[13] A. Ya. Khinchin. *Continued Fractions*. The University of Chicago Press, 1964.

[14] M. Lothaire. *Combinatorics on Words*. Cambridge Academic Press, second edition, 1997.

[15] M. Lothaire. *Algebraic Combinatorics on Words*. Cambridge Academic Press, 2002.

[16] M. Morse and G.A. Hedlund. Symbolic dynamics II : Sturmian trajectories. *Amer. J. Math.*, 62:1–42, 1940.

[17] Dinard van der Laan. *The structure and performance of optimal routing sequences*. PhD thesis, Leiden University, 2003.

Unité de recherche INRIA Rhône-Alpes
655, avenue de l'Europe - 38334 Montbonnot Saint-Ismier (France)

Unité de recherche INRIA Futurs : Parc Club Orsay Université - ZAC des Vignes
4, rue Jacques Monod - 91893 ORSAY Cedex (France)
Unité de recherche INRIA Lorraine : LORIA, Technopôle de Nancy-Brabois - Campus scientifi que
615, rue du Jardin Botanique - BP 101 - 54602 Villers-lès-Nancy Cedex (France)
Unité de recherche INRIA Rennes : IRISA, Campus universitaire de Beaulieu - 35042 Rennes Cedex (France)
Unité de recherche INRIA Rocquencourt : Domaine de Voluceau - Rocquencourt - BP 105 - 78153 Le Chesnay Cedex (France)
Unité de recherche INRIA Sophia Antipolis : 2004, route des Lucioles - BP 93 - 06902 Sophia Antipolis Cedex (France)

Éditeur
INRIA - Domaine de Voluceau - Rocquencourt, BP 105 - 78153 Le Chesnay Cedex (France)
http://www.inria.fr
ISSN 0249-6399